

WO 92/10092

PCT/US91/08693

57

location and the process is repeated. In preferred embodiments the data are acquired every 1 to 100 μm with a data collection diameter of about 0.8 to 10 μm preferred. In embodiments with sufficiently high fluorescence, a CCD detector with broadfield illumination is utilized.

Fig. 11 illustrates the architecture of the data collection system in greater detail. Operation of the system occurs under the direction of the photon counting program 1102. The user inputs the scan dimensions, the number of pixels or data points in a region, and the scan speed to the counting program. Via a GPIB bus 1104 the program (in an IBM PC compatible computer, for example) interfaces with a multichannel scaler 1106 such as a Stanford Research SR 430 and an x-y stage controller 1108 such as a Newport PM500. The signal from the light from the fluorescing substrate enters a photomultiplier 1110, providing output to the scaler 1106. Data are output from the scaler indicative of the number of counts in a given region. After scanning a selected area, the stage controller is activated with commands for acceleration and velocity, which in turn drives the scan stage 1112 such as a Newport PM500-A to another region.

Data are collected in an image data file 1114 and processed in a scaling program 1116. A scaled image is output for display on, for example, a VGA display 1118. The image is scaled based on an input of the percentage of pixels to clip and the minimum and maximum pixel levels to be viewed. The system outputs for use the min and max pixel levels in the raw data.

B. Data Analysis

The output from the data collection system is an array of data indicative of fluorescence intensity versus location on the substrate. The data are typically taken over regions substantially smaller than the area in

WO 92/10092

PCT/US91/08693

58

which synthesis of a given polymer has taken place. Merely by way of example, if polymers were synthesized in squares on the substrate having dimensions of 500 microns by 500 microns, the data may be taken over regions having dimensions of 5 microns by 5 microns. In most preferred embodiments, the regions over which fluorescence data are taken across the substrate are less than about $1/2$ the area of the regions in which individual polymers are synthesized, preferably less than $1/10$ the area in which a single polymer is synthesized, and most preferably less than $1/100$ the area in which a single polymer is synthesized. Hence, within any area in which a given polymer has been synthesized, a large number of fluorescence data points are collected.

A plot of the number of pixels versus fluorescence intensity for a scan of a cell when it has been exposed to, for example, a labeled antibody will typically take the form of a bell curve, but spurious data are observed, particularly at higher intensities. Since it is desirable to use an average of fluorescence intensity over a given synthesis region in determining relative binding affinity, these spurious data will tend to undesirably skew the data.

Accordingly, in one embodiment of the invention the data are corrected for removal of these spurious data points, and an average of the data points is thereafter utilized in determining relative binding efficiency.

Fig. 12 illustrates one embodiment of a system for removal of spurious data from a set of fluorescence data such as data used in affinity screening studies. A user or the system inputs data relating to the chip location and cell corners at step 1302. From this information and the image file, the system creates a computer representation of a histogram at step 1304, the histogram (at least in the form of a computer file) plotting number of data pixels versus intensity.

WO 92/10092

PCT/US91/08693

59

For each cell, a main data analysis loop is then performed. For each cell, at step 1306, the system calculates the total fluorescence intensity or number of pixels for the bandwidth centered around varying intensity levels. For example, as shown in the plot to the right of step 1306, the system calculates the number of pixels within the band of width w . The system then "moves" this bandwidth to a higher center intensity, and again calculates the number of pixels in the bandwidth. This process is repeated until the entire range of intensities have been scanned, and at step 1308 the system determines which band has the highest total number of pixels. The data within this bandwidth are used for further analysis. Assuming the bandwidth is selected to be reasonably small, this procedure will have the effect of eliminating spurious data located at the higher intensity levels. The system then repeats at step 1310 if all cells have been evaluated, or repeats for the next cell.

At step 1312 the system then integrates the data within the bandwidth for each of the selected cells, sorts the data at step 1314 using the synthesis procedure file, and displays the data to a user on, for example, a video display or a printer.

V. Representative Applications

A. Oligonucleotide Synthesis

The generality of light directed spatially addressable parallel chemical synthesis is demonstrated by application to nucleic acid synthesis.

1. Example

Light activated formation of a thymidine-cytidine dimer was carried out. A three dimensional representation of a fluorescence scan showing a 7 square by 4 square checkerboard pattern generated by the light-directed synthesis of a dinucleotide was produced.

WO 92/10092

PCT/US91/08693

60

5'-nitroveratryl thymidine was attached to a synthesis substrate through the 3' hydroxyl group. The nitroveratryl protecting groups were removed by illumination through a 500 μm checkerboard mask. The substrate was then treated with phosphoramidite activated 2'-deoxycytidine. In order to follow the reaction fluorometrically, the deoxycytidine had been modified with an Fmoc protected aminohexyl linker attached to the exocyclic amine (5'-O-dimethoxytrityl-4-N-(6-N-fluorenylmethylcarbamoyl-hexylcarboxy)-2'-deoxycytidine). After removal of the Fmoc protecting group with base, the regions which contained the dinucleotide were fluorescently labelled by treatment of the substrate with 1 mM FITC in DMF for one hour.

The three-dimensional representation of the fluorescence intensity data showing alternating squares of bright raised pixels reproduces the checkerboard illumination pattern used during photolysis of the substrate. This result demonstrates that oligonucleotides as well as peptides can be synthesized by the light-directed method.

In another example the light-activated formation of thymidine-cytidine-cytidine was carried out as shown in Fig. 13. Here, as in the previous example, 5'-nitroveratryl thymidine was attached to the substrate, via phosphoramidite chemistry to a surface containing [Bis (2-hydroxyethyl)-3-aminopropylsiloxane]. The slide was then uniformly illuminated (362nm at $\sim 14\text{mW}/\text{cm}^2$) for 10 minutes in the presence of dioxane. After drying, the surface was then treated with N,4-dimethoxytrityl-5'-nitroveratryl-2'-deoxycytidine-3'-O-(2-cyanoethyl)-N,N-diisopropylphosphoramidite in the presence of tetrazole (standard phosphoramidite coupling chemistry). After oxidizing and drying, the plate was again illuminated as before except that a 500 μm checkerboard mask was placed between the light source and the slide. The surface was then exposed to 5'-O-(4,4'-Dimethoxy)-N-4-(6-

WO 92/10092

PCT/US91/08693

61

((Biotinoyl)amino)hexanoyl)amino)hexanoyl, aminohexyl)-5-methyl-2'-deoxycytidine-3'-O-(2-cyanoethyl)-N,N-diisopropylphosphoramidite with tetrazale. After oxidizing and drying, the areas which contained the trinucleotide were fluoroesciently labelled by treatment with FITC labled streptavidin. A resulting representation of the fluorescence intensity data showed alternating bright and dark squares corresponding to the 500 μ m and checkerboard illumination pattern used during photolysis.

VI. Conclusion

The inventions herein provide a new approach for the simultaneous synthesis of a large number of compounds. The method can be applied whenever one has chemical building blocks that can be coupled in a solid-phase format, and when light can be used to generate a reactive group.

The above description is illustrative and not restrictive. Many variations of the invention will become apparent to those of skill in the art upon review of this disclosure. Merely by way of example, while the invention is illustrated primarily with regard to peptide and nucleotide synthesis, the invention is not so limited. The scope of the invention should, therefore, be determined not with reference to the above description, but instead should be determined with reference to the appended claims along with their full scope of equivalents.

WO 92/10092

PCT/US91/08693

62

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Tyr Gly Gly Phe Leu
1 5

WO 92/10092

63

PCT/US91/08693

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Pro Gly Gly Phe Leu
1 5

WO 92/10092

PCT/US91/08693

64

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Tyr Gly Ala Gly Phe
1 5

WO 92/10092

PCT/US91/08693

65

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 6 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Tyr Gly Ala Phe Leu Ser
1 5

WO 92/10092

PCT/US91/08693

66

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Tyr	Gly	Ala	Phe	Ser
1				5

WO 92/10092

PCT/US91/08693

67

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Tyr Gly Ala Phe Leu
1 5

WO 92/10092

PCT/US91/08693

68

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 6 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Tyr Gly Gly Phe Leu Ser
1 5

WO 92/10092

69

PCT/US91/08693

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Tyr Gly Ala Phe
1

WO 92/10092

PCT/US91/08693

70

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Tyr	Gly	Ala	Leu	Ser
1				5

WO 92/10092

PCT/US91/08693

71

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Tyr Gly Gly Phe Ser
1 5

WO 92/10092

72

PCT/US91/08693

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 4 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Tyr Gly Ala Leu
1

WO 92/10092

73

PCT/US91/08693

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 6 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Tyr Gly Ala Phe Leu Phe
1 5

WO 92/10092

74

PCT/US91/08693

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Tyr Gly Ala Phe Phe
1 5

WO 92/10092

PCT/US91/08693

75

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Tyr Gly Gly Leu Ser
1 5

WO 92/10092

PCT/US91/08693

76

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Tyr Gly Gly Phe Leu
1 5

WO 92/10092

PCT/US91/08693

77

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 6 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Tyr Gly Ala Phe Ser Phe
1 5

WO 92/10092

PCT/US91/08693

78

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 7 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Tyr Gly Ala Phe Leu Ser Phe
1 5

WO 92/10092

79

PCT/US91/08693

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 6 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Tyr	Gly	Ala	Phe	Met	Gln
1				5	

WO 92/10092

PCT/US91/08693

80

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Tyr Gly Ala Phe Met
1 5

WO 92/10092

PCT/US91/08693

81

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Tyr Gly Ala Phe Gln
1 5

WO 92/10092

PCT/US91/08693

82

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Tyr Gly Gly Phe Met
1 5

WO 92/10092

PCT/US91/08693

83

WHAT IS CLAIMED IS:

1. A reactor system for synthesizing a plurality of polymer sequences on a substrate comprising:
 - a) a reactor for contacting reaction fluids to said substrate;
 - b) a system for delivering selected reaction fluids to said reactor;
 - c) a translation stage for moving a mask or substrate from at least a first relative location relative to a second relative location;
 - d) a light for illuminating said substrate through a mask at selected times; and
 - e) an appropriately programmed digital computer for selectively directing a flow of fluids from said reactor system, selectively activating said translation stage, and selectively illuminating said substrate so as to form a plurality of diverse polymer sequences on said substrate at predetermined locations.
2. The reactor system as recited in claim 1 adapted to provide a plurality of monomers in a reaction fluid to said substrate, said substrate used for an initial screening of polymer sequences.
3. An ordered method for forming a plurality of polymer sequences by sequential addition of reagents comprising the step of serially protecting and deprotecting portions of said plurality of polymer sequences for addition of other portions of said polymer sequences using a binary synthesis strategy.
4. The method as recited in claim 3 wherein said binary synthesis strategy is a binary masking strategy.
5. The method as recited in claim 4 wherein said masking strategy in which said masking strategy provides

WO 92/10092

PCT/US91/08693

84

at least two consecutive steps in which a mask factors a previous mask by protecting a portion of a previously illuminated portions to light and exposing a portion of a previously protected portions to light.

6. The method as recited in claim 4 in which said masking strategy in which at least two successive steps in said masking strategy illuminate about one half of a region of interest on said substrate.

7. The method as recited in claim 4 wherein said masking strategy forms a plurality of polymer sequences on a single substrate.

8. The method as recited in claim 4 wherein said masks are arranged in a gray code masking strategy, said gray code masking strategy having one edge illumination on each of a plurality of synthesis sites.

9. The method as recited in claim 4 wherein said masking strategy results in a minimum number of masking steps for a number of polymers synthesized.

10. The method as recited in claim 4 wherein all possible polymers of length 1 are formed with a given basis set of monomers.

11. The method as recited in claim 4 wherein said masking strategy is developed in an appropriately programmed digital computer inputting at least a desired basis set, and length of polymers.

12. The method as recited in claim 4 wherein all possible polymers of a length less than or equal to 1 are formed with a given basis set of monomers.

WO 92/10092

PCT/US91/08693

85

13. The method as recited in claim 4 further comprising the step of forming a portion of said polymers with a non-binary masking strategy.

14. The method as recited in claim 10 further comprising the step of outputting a masking strategy.

15. The method as recited in claim 10 further comprising the step of outputting a map of synthesized polymers on said substrate.

16. The method as recited in claim 15 wherein said map is in the form of Fig. 9.

17. A method of screening a plurality of linker polymers for use in binding affinity studies comprising the steps of:

a) forming a plurality of linker polymers on a substrate in selected regions, said linker polymers formed by the steps of recursively:

i) on a surface of a substrate, irradiating a portion of said selected regions to remove a protecting group; and

ii) contacting said surface with a monomer;

b) contacting said plurality of linker polymers with a ligand; and

c) contacting said ligand with a labeled receptor.

18. The method as recited in claim 17 wherein said ligand is a polypeptide.

19. The method as recited in claim 17 wherein said receptor is an antibody.

20. The method as recited in claim 17 wherein said monomers added in step ii) are the same in each of said

WO 92/10092

PCT/US91/08693

86

recursive steps, said selected regions comprising linker molecules of different lengths.

21. The method as recited in claim 17 wherein said labelled receptor is a fluoresceinated receptor.

22. A system for determining affinity of a receptor to a ligand comprising:

a) means for applying light to a surface of a substrate, said substrate comprising a plurality of ligands at predetermined locations, said means for applying directing light providing simultaneous illumination at a plurality of said predetermined locations; and

b) an array of detectors for detecting fluorescence at said plurality of predetermined locations.

23. A system as recited in claim 22 wherein said means for applying light comprises a point light source and a cylindrical lens for focusing said point light source along a substantially linear path.

24. A system as recited in claim 22 wherein said array of detectors comprises a linear array.

25. A system as recited in claim 22 wherein said array of detectors comprises a linear CCD array.

26. In a digital computer, a method of determining the tendency of a receptor to bind to a ligand comprising:

a) exposing fluorescently labelled receptors to a substrate, said substrate comprising a plurality of ligands in regions at known locations;

WO 92/10092

PCT/US91/08693

87

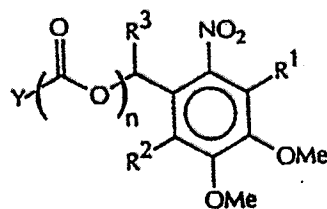
b) at a plurality of data collection points within each of said regions, determining an amount of fluorescence from said data collection points;

c) removing said data collection points deviating from a preset amount from a predetermined statistical distribution; and

d) determining a relative binding affinity of said receptor to remaining data collection points.

27. The method as recited in claim 26 wherein said predetermined statistical distribution is a normal distribution.

28. A compound having the formula:



wherein $n = 0$ or 1 ; Y is selected from the group consisting of an oxygen of the carboxyl group of a natural or unnatural amino acid, an amino group of a natural or unnatural amino acid, or the C-5' oxygen group of a natural or unnatural deoxyribonucleic or ribonucleic acid; R^1 and R^2 independently are a hydrogen atom, a lower alkyl, aryl, benzyl, halogen, hydroxyl, alkoxyl, thiol, thioether, amino, nitro, carboxyl, formate, formamido, sulfido, or phosphido group; and R^3 is a alkoxy, alkyl, aryl, hydrogen, or alkenyl group.

29. The compound of claim 28 wherein Y is the C-5' oxygen group of a natural or unnatural deoxyribonucleic or ribonucleic acid.

30. The compound of claim 29 wherein $n = 0$.

WO 92/10092

PCT/US91/08693

88

31. The compound of claim 29 wherein R^1 and R^2 are each a hydrogen atom.

32. The compound of claim 31 wherein R^3 is a hydrogen atom.

33. The compound of claim 31 wherein R^3 is a methyl group.

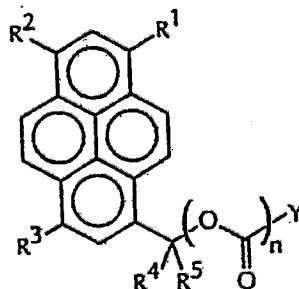
34. The compound of claim 28 wherein Y is an oxygen of the carboxyl group of an amino acid and $n = 0$.

35. The compound of claim 34 wherein R^1 and R^2 are each a hydrogen atom.

36. The compound of claim 35 wherein R^3 is a hydrogen atom.

37. The compound of claim 35 wherein R^3 is a methyl group.

38. A compound having the formula:



wherein $n = 0$ or 1 ; Y is selected from the group consisting of an amino group of a natural or unnatural amino acid or the C-5' oxygen group of a natural or unnatural deoxyribonucleic and ribonucleic acid; R^1 , R^2 , and R^3 independently are a hydrogen atom, a lower alkyl,

WO 92/10092

PCT/US91/08693

89

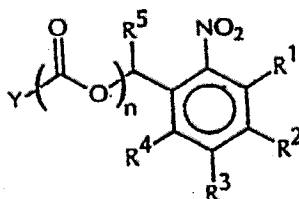
aryl, benzyl, halogen, hydroxyl, alkoxyl, thiol, thioether, amino, nitro, carboxyl, formate, formamido, sulfido or phosphido group; R^4 and R^5 independently are a alkoxy, alkyl, hydrogen, halo, aryl, or alkenyl group.

39. The compound of claim 38 wherein R^1 through R^3 are each a hydrogen atom.

40. The compound of claim 39 wherein R^4 and R^5 are each a hydrogen atom.

41. The compound of claim 39 wherein R^4 and R^5 are each a methyl group.

42. A compound having the formula:



wherein $n = 0$ or 1 ; Y is a C-5' oxygen group of a natural or unnatural deoxyribonucleic and ribonucleic acid; R^1 through R^4 independently are a hydrogen atom, a lower alkyl, aryl, benzyl, halogen, hydroxyl, alkoxyl, thiol, thioether, amino, nitro, carboxyl, formate, formamido, sulfido, or phosphido group; and R^5 is a alkoxy, alkyl, aryl, or alkenyl group.

43. The compound of claim 42 wherein R^2 and R^3 are each a methoxy group.

44. The compound of claim 43 wherein R^1 and R^4 are each a hydrogen atom.

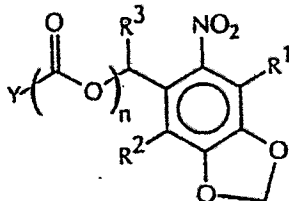
WO 92/10092

PCT/US91/08693

90

45. The compound of claim 44 wherein R^5 is a methyl group.

46. A compound having the formula:



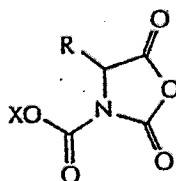
wherein $n = 0$ or 1 ; Y is an atom to be protected; R^1 and R^2 independently are a hydrogen atom, a lower alkyl, aryl, benzyl, halogen, hydroxyl, alkoxyl, thiol, thioether, amino, nitro, carboxyl, formate, formamido, sulfido, or phosphido group; and R^3 is a alkoxy, alkyl, aryl, or alkenyl group.

47. The compound of claim 46 wherein Y is selected from the group consisting of an oxygen of the carboxyl group of a natural or unnatural amino acid, or the C-5' oxygen group of a natural or unnatural deoxyribonucleic or ribonucleic acid, or the amino group of a natural or unnatural amino acid.

48. The compound of claim 47 wherein R^1 and R^2 are hydrogen.

49. The compound of claim 48 wherein R^3 is a methyl group.

50. A compound having the formula:



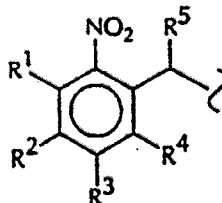
WO 92/10092

PCT/US91/08693

91

where R is a side chain of a natural or unnatural amino acid and X is a photoremovable protecting group.

51. The compound of claim 50 wherein X has the following formula:



where R¹, R², R³, and R⁴ independently are a hydrogen atom, a lower alkyl, aryl, benzyl, halogen, hydroxyl, alkoxyl, thiol, thioether, amino, nitro, carboxyl, formate, formamido or phosphido group, or adjacent substituents are substituted oxygen groups that together form a cyclic acetal or ketal; and R⁵ is a hydrogen atom, a alkoxyl, alkyl, halo, aryl, or alkenyl group.

52. The compound of claim 51 wherein R¹ and R⁴ are each a hydrogen atom, and R² and R³ are each a methoxy group.

53. The compound of claim 52 wherein R⁵ is a methyl group.

54. The compound of claim 51 wherein R² and R³ are substituted oxygen groups that together form a cyclic acetal.

55. The compound of claim 54 wherein R¹ and R⁴ are each a hydrogen atom.

WO 92/10092

PCT/US91/08693

92

56. The compound of claim 55 wherein R⁵ is a methyl group.

WO 92/10092

1/12

PCT/US91/08693

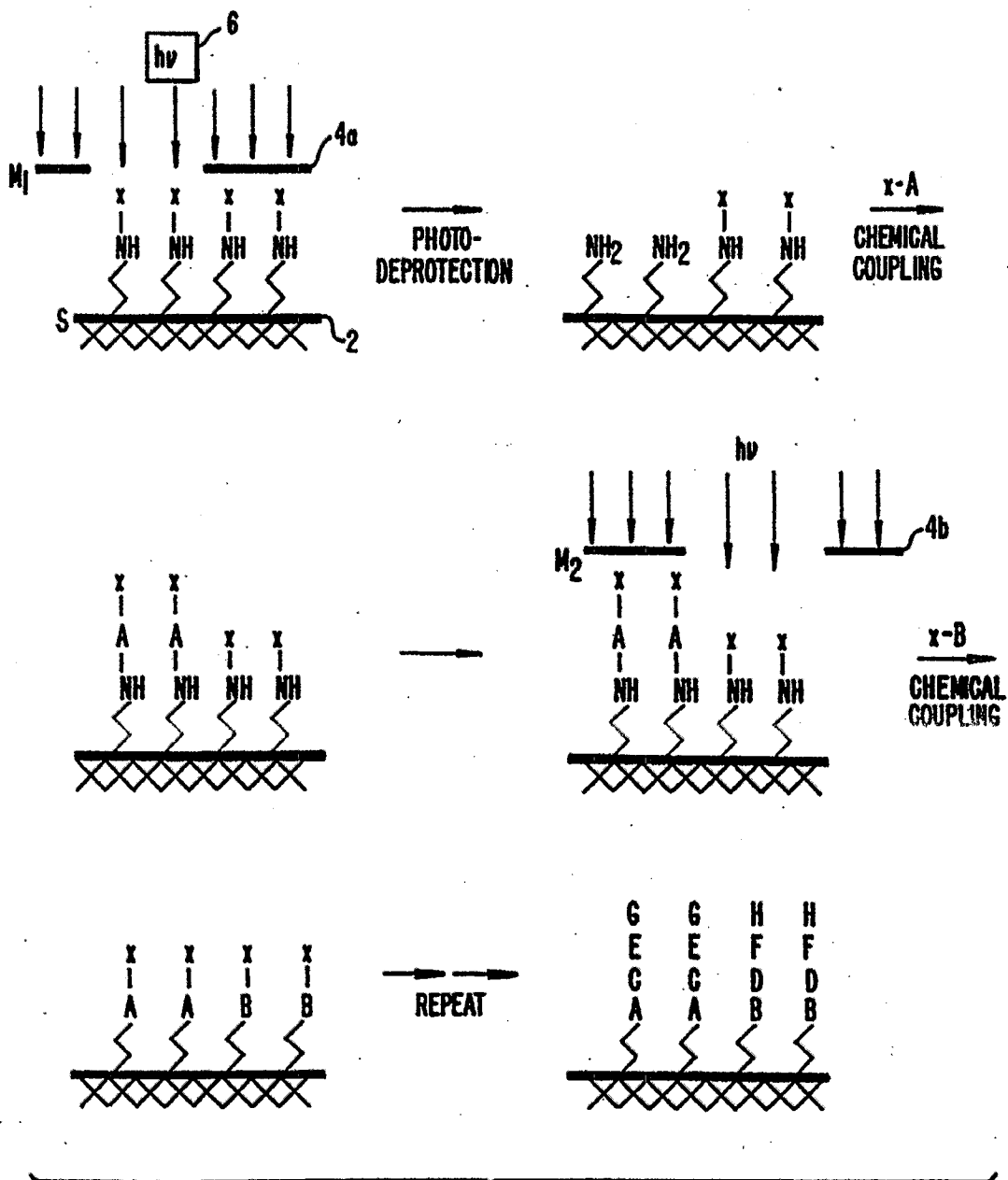


FIG. 1.

SUBSTITUTE SHEET

WO 92/10092

2 / 16

PCT/US91/08693

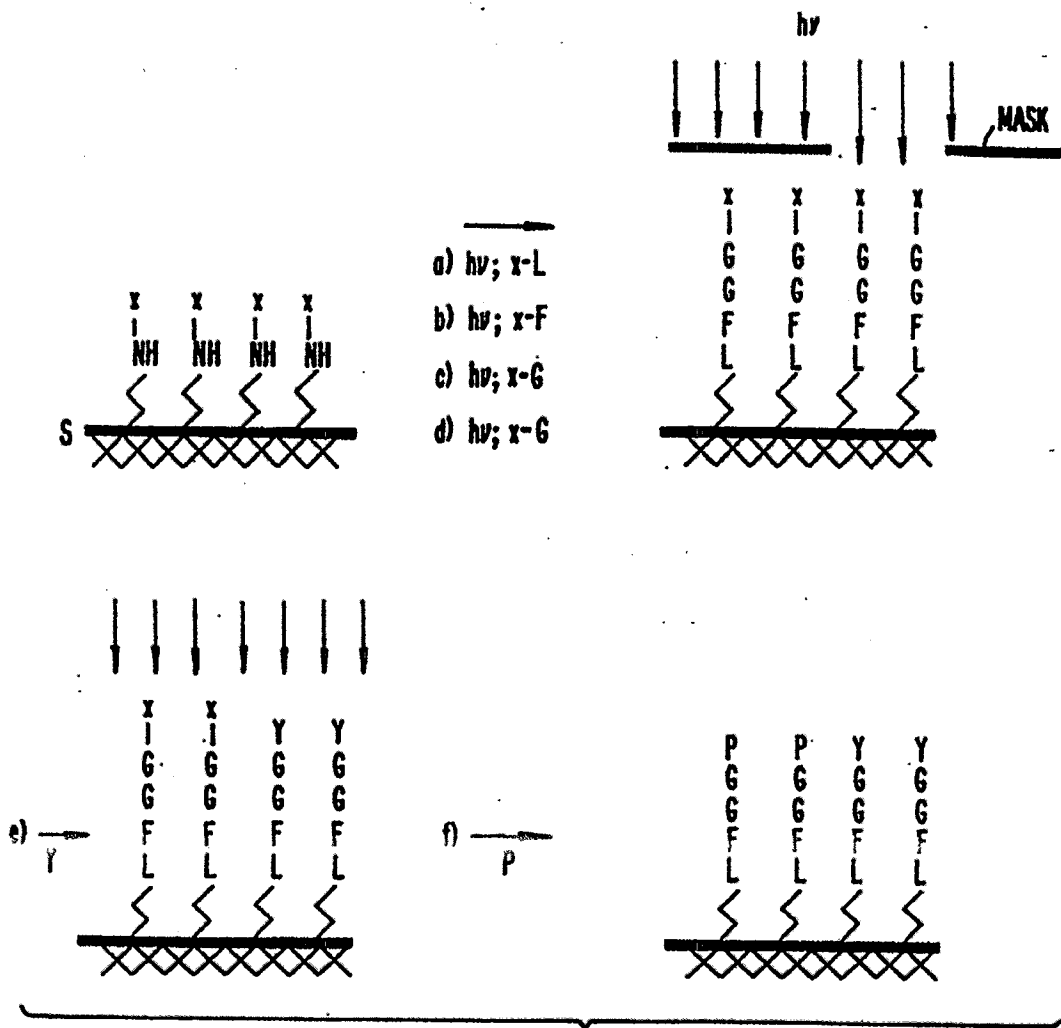


FIG. 2.

WO 92/10092

PCT/US91/08693

3 / 16

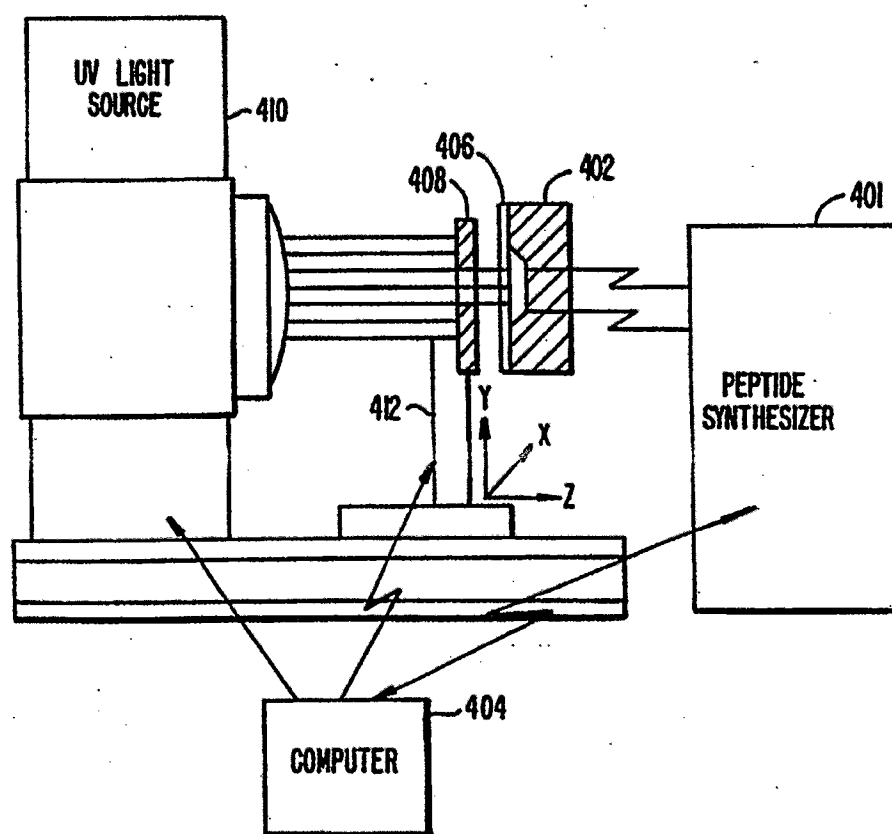


FIG. 3.

SUBSTITUTE SHEET

WO 92/10092

4/16

PCT/US91/08693

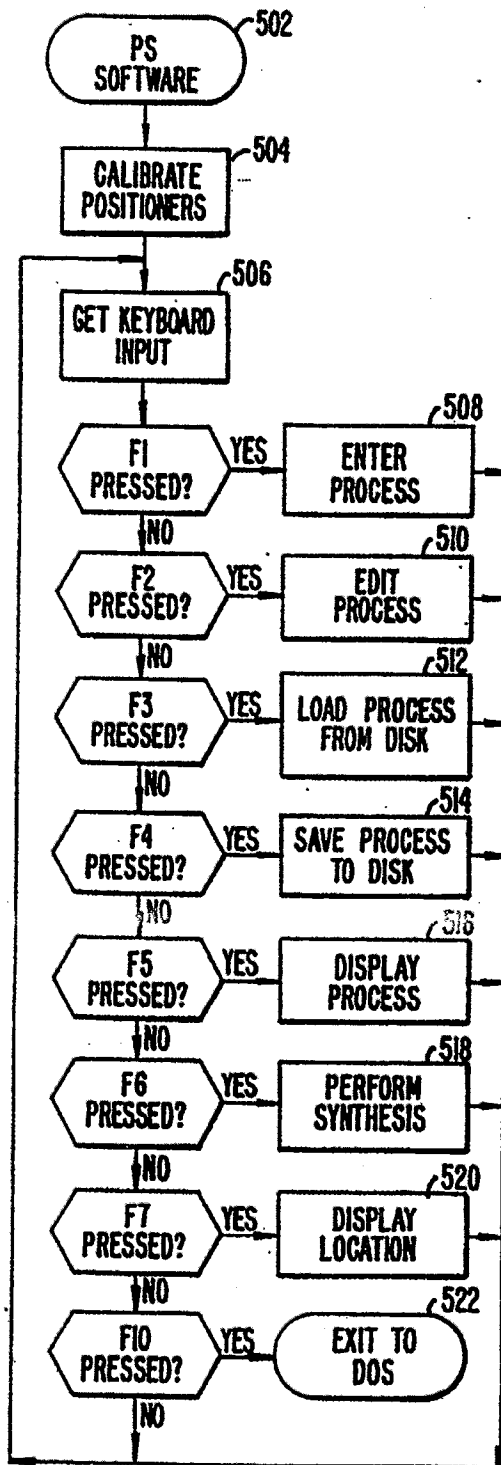


FIG. 4A.

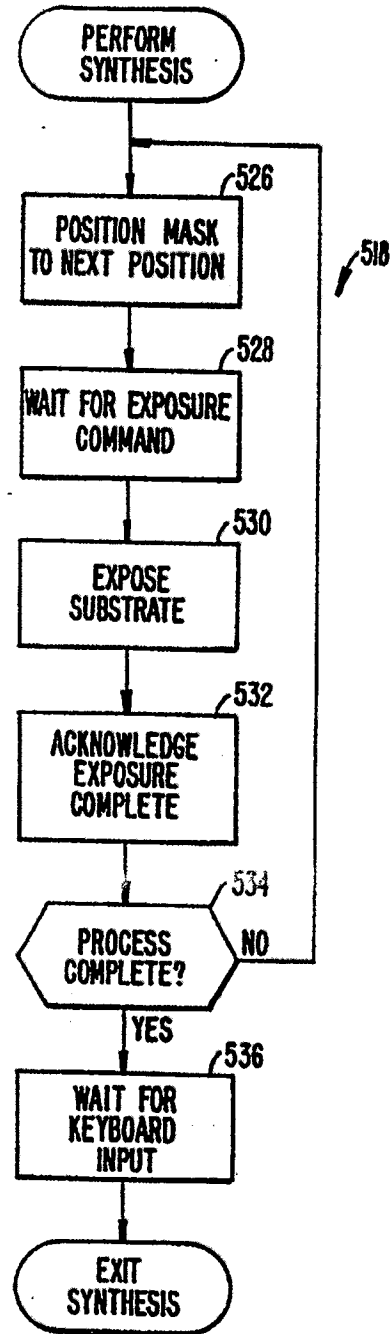


FIG. 4B.

WO 92/10092

5/16

PCT/US91/08693

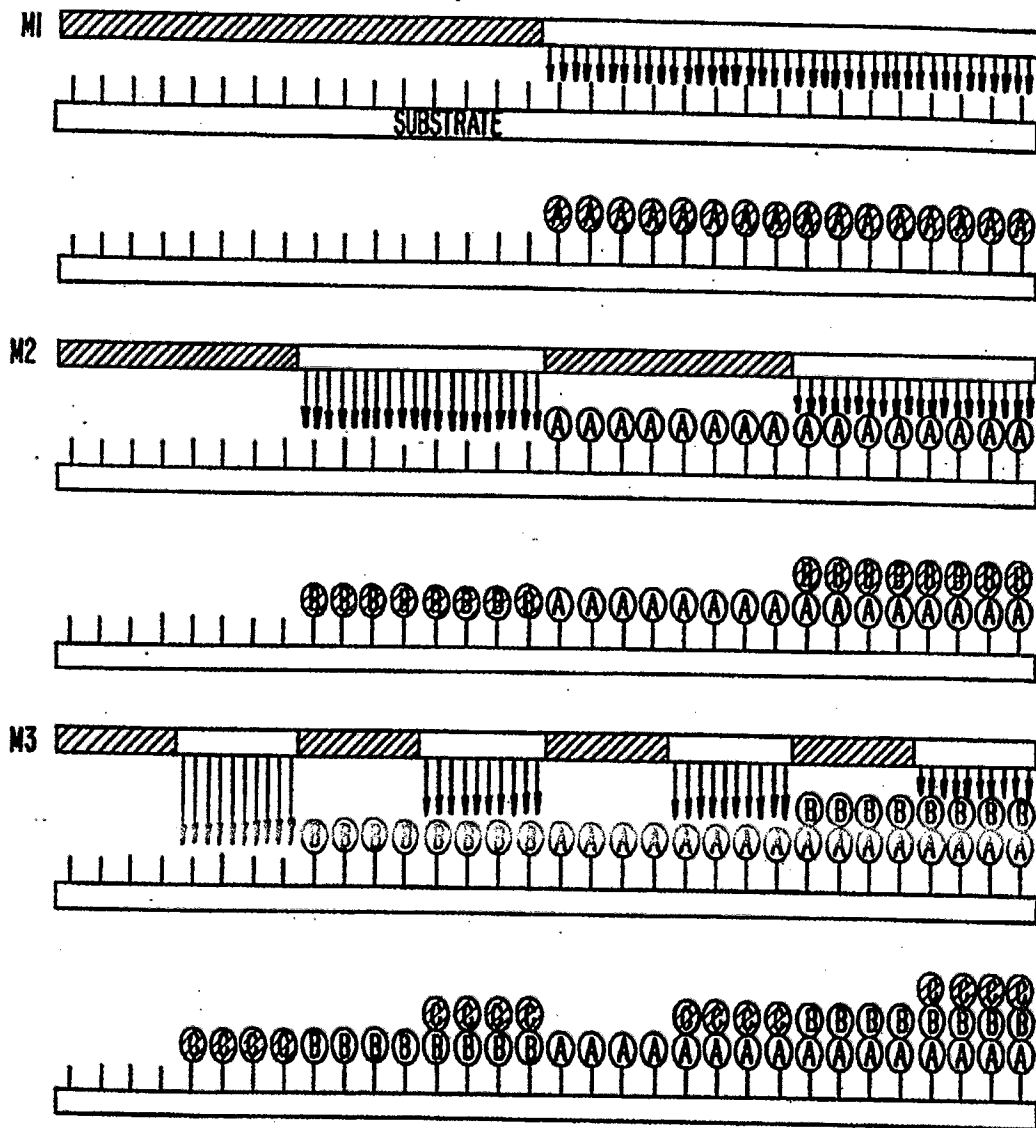


FIG. 5A.

FIG. 5A.
FIG. 5B.

FIG. 5.

WO 92/10092

6 / 16

PCT/US91/08693

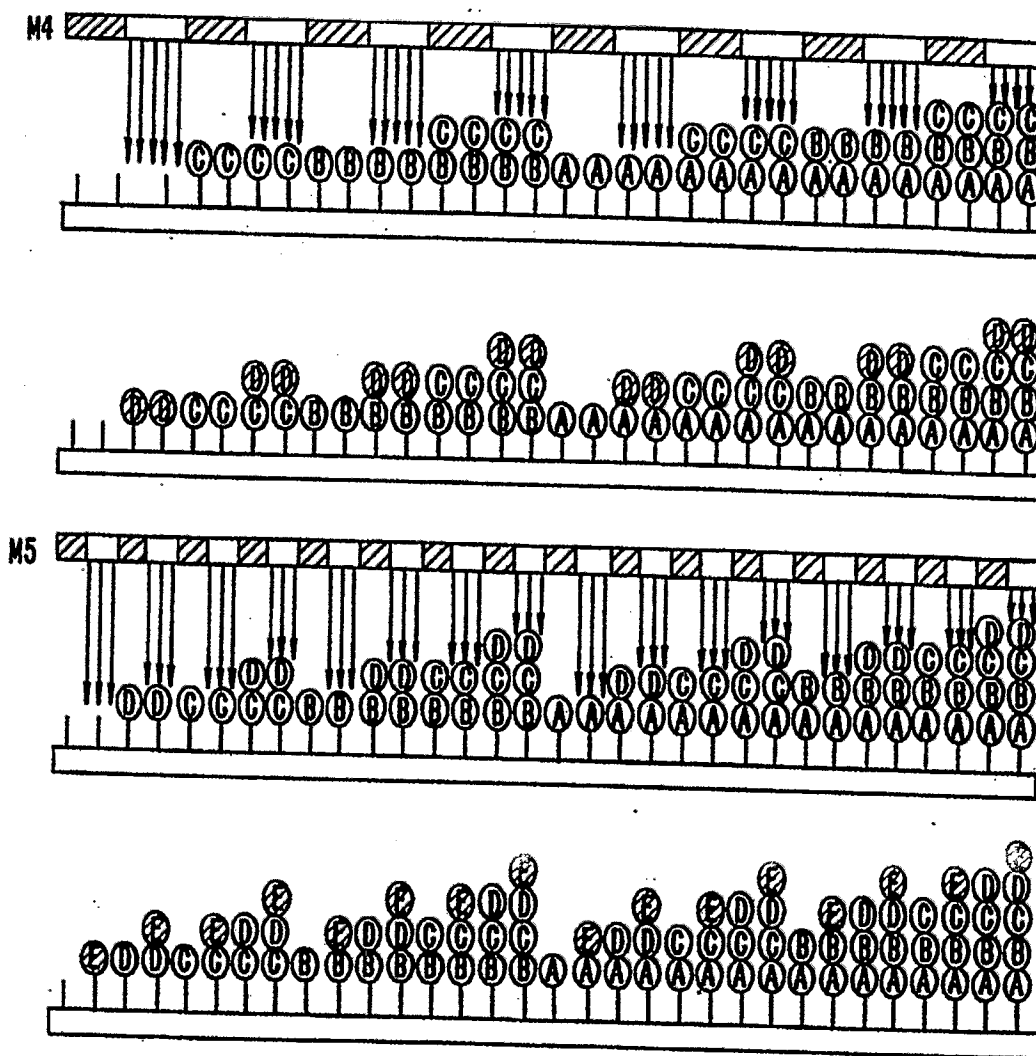


FIG. 5B.

WO 92/10092

7/16

PCT/US91/08693

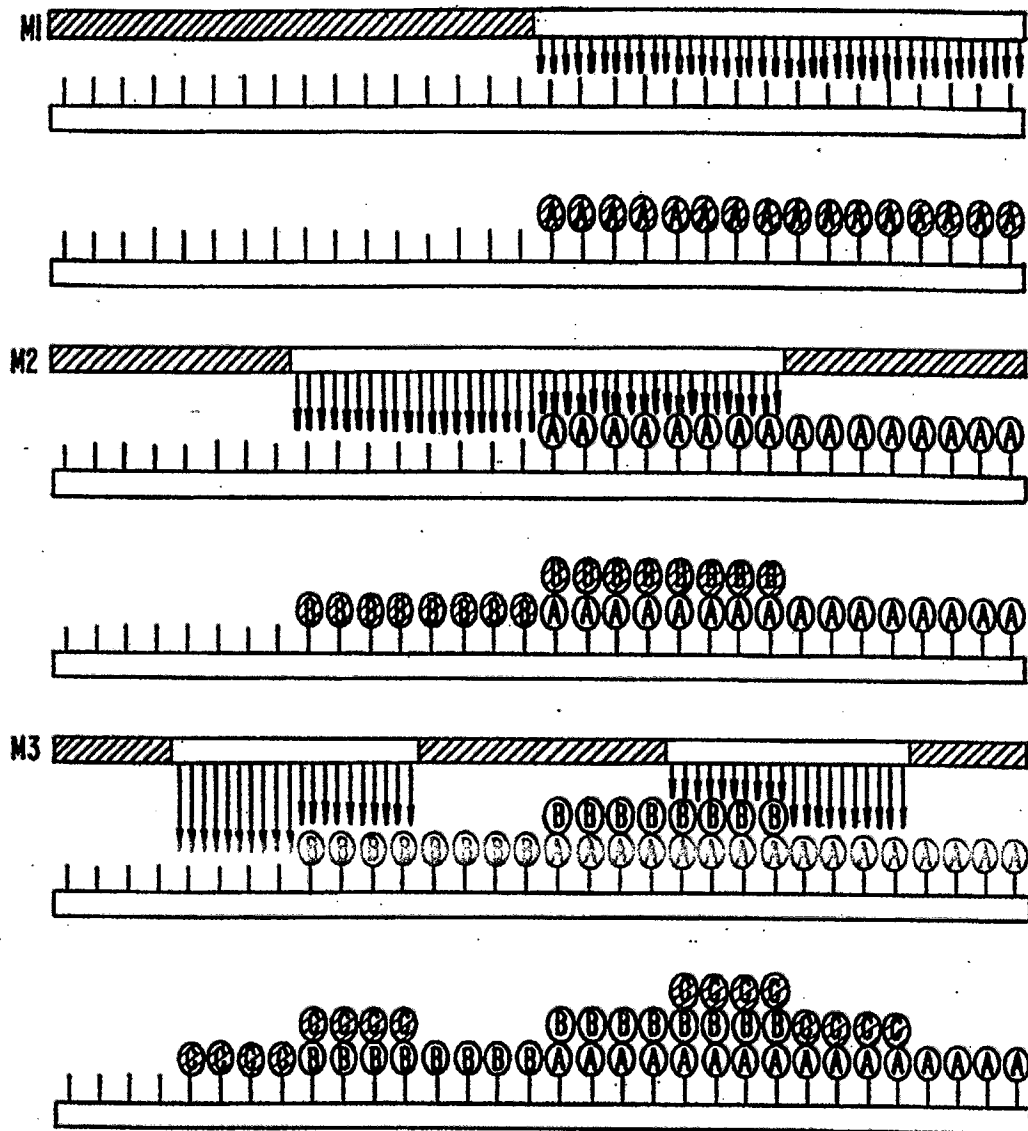


FIG. 6A.

FIG. 6A.
FIG. 6B.

FIG. 6.

SUBSTITUTE SHEET

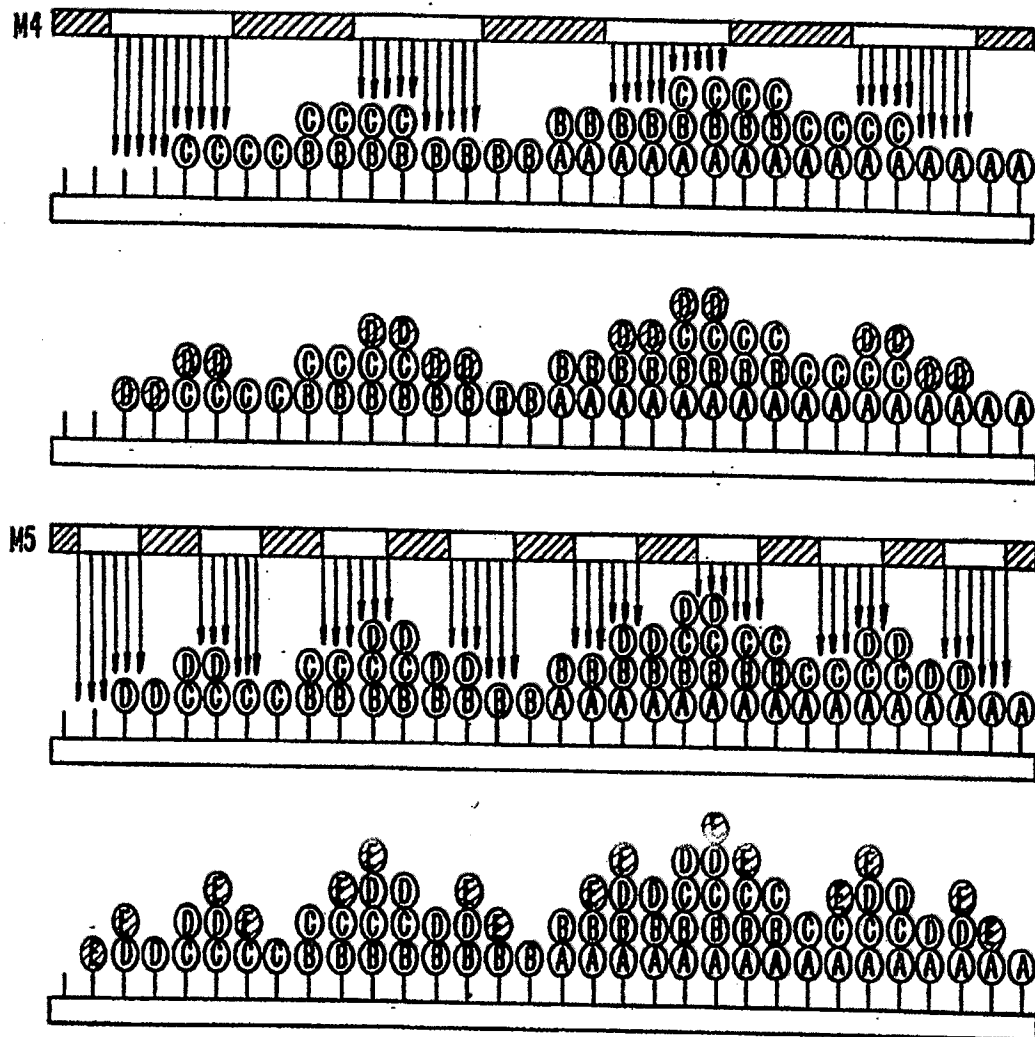


FIG. 6B.

SUBSTITUTE SHEET

WO 92/10092

9 / 16

PCT/US91/08693

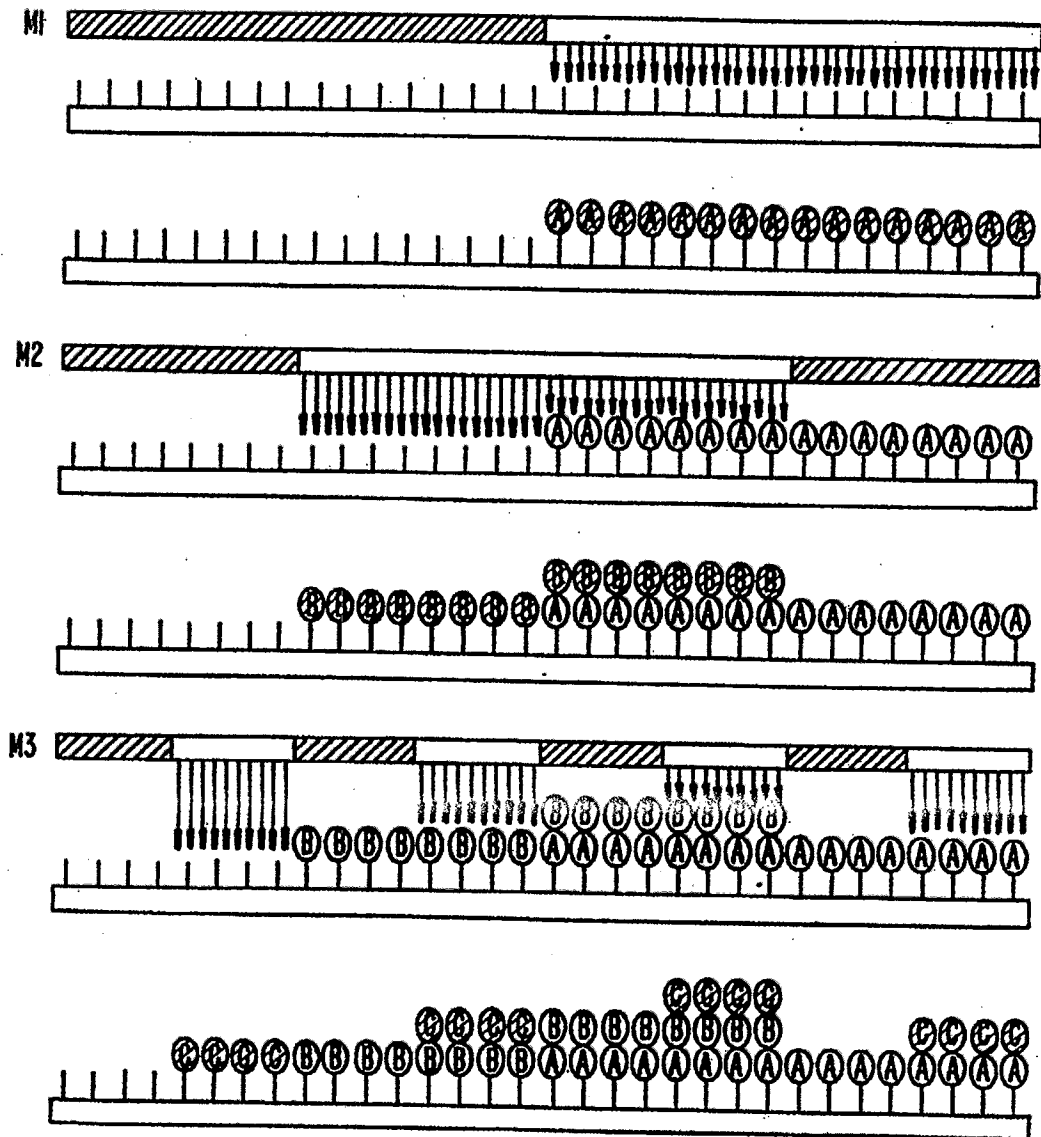


FIG. 7A.

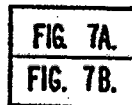


FIG. 7.

SUBSTITUTE SHEET

WO 92/10092

10 / 16

PCT/US91/08693

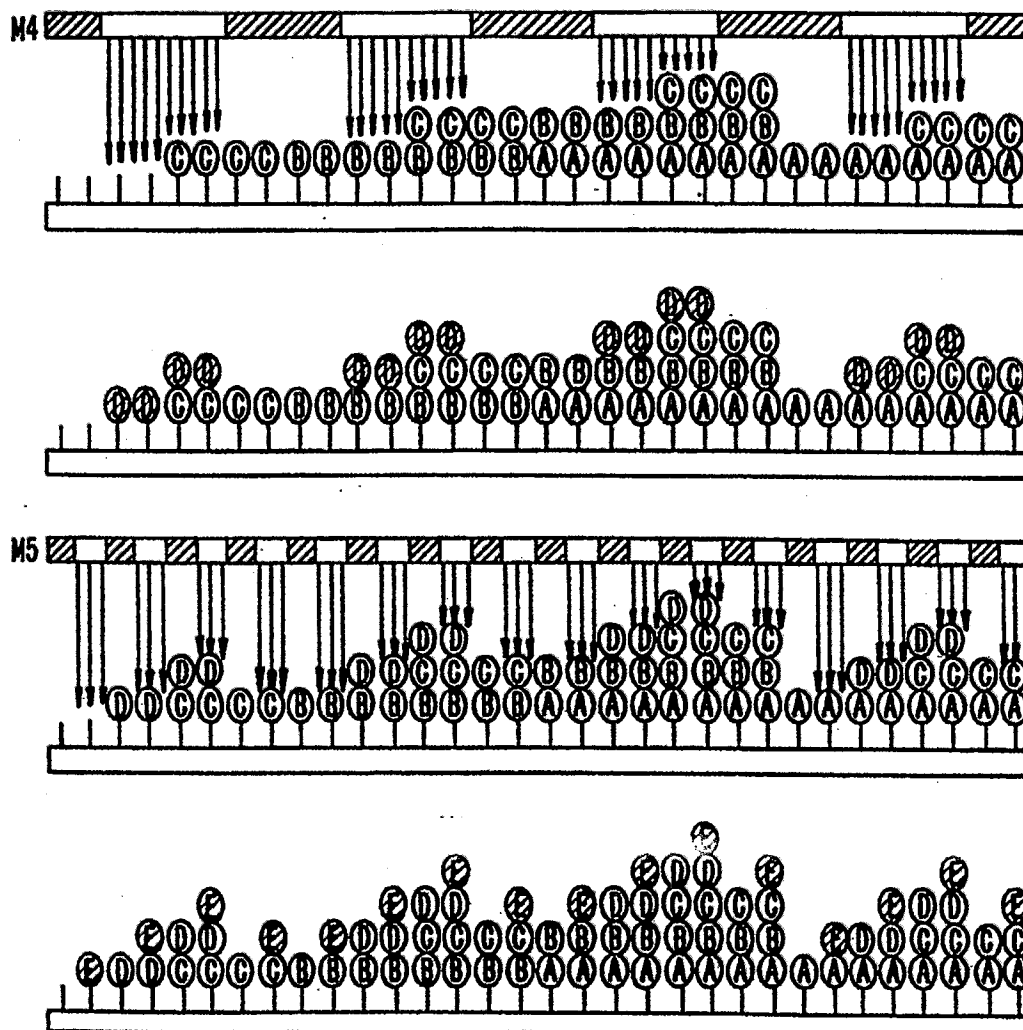


FIG. 7B.

SUBSTITUTE SHEET

WO 92/10092

11 / 16

PCT/US91/08693

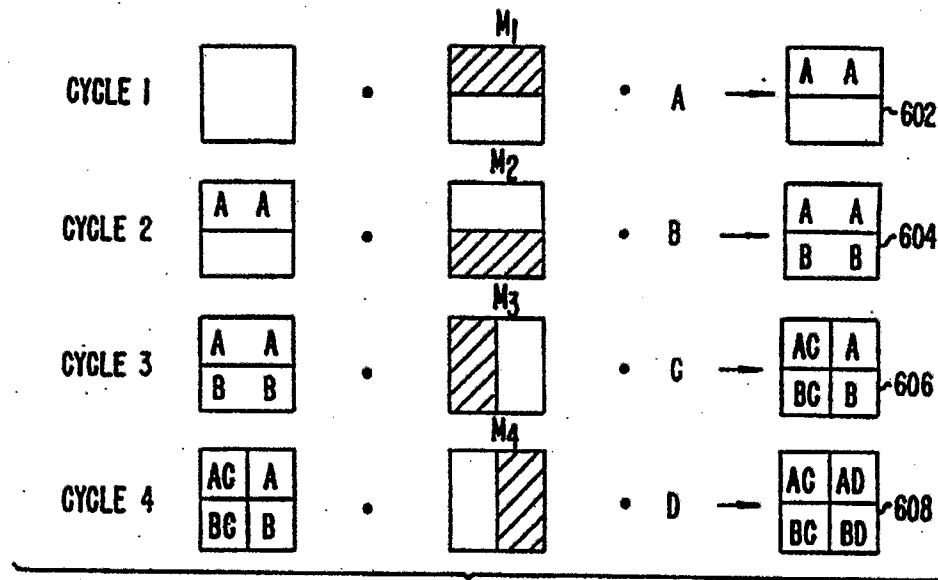


FIG. 8A.

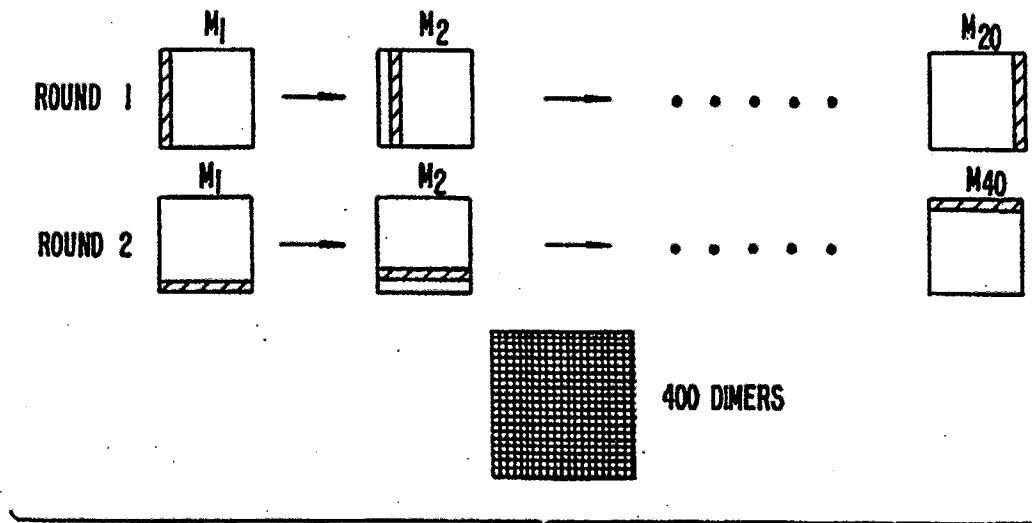


FIG. 8B.

SUBSTITUTE SHEET

WO 92/10092

12 / 16

PCT/US91/08693

[illegible]

FIG. 9.

SUBSTITUTE SHEET

WO 92/10092

13 / 16

PCT/US91/08693

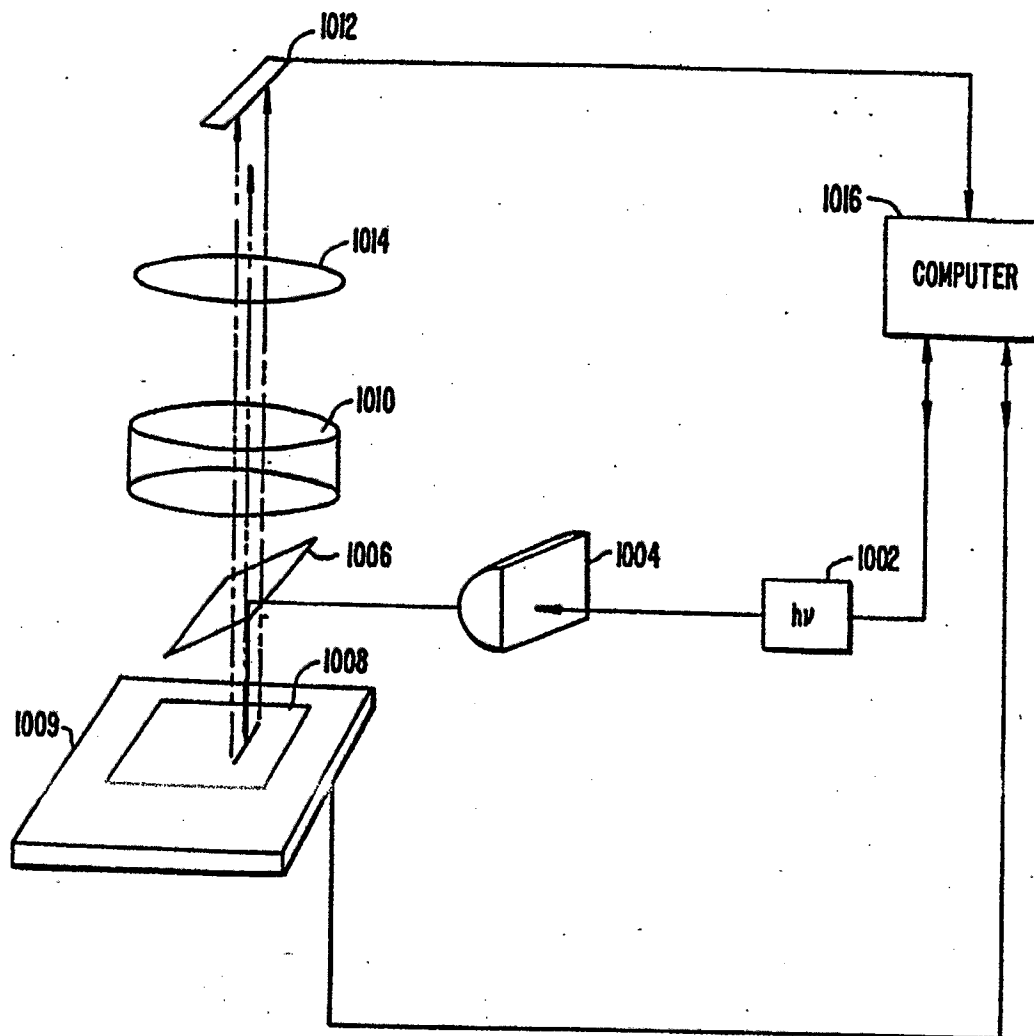


FIG. 10.

SUBSTITUTE SHEET

WO 92/10092

14 / 16

PCT/US91/08693

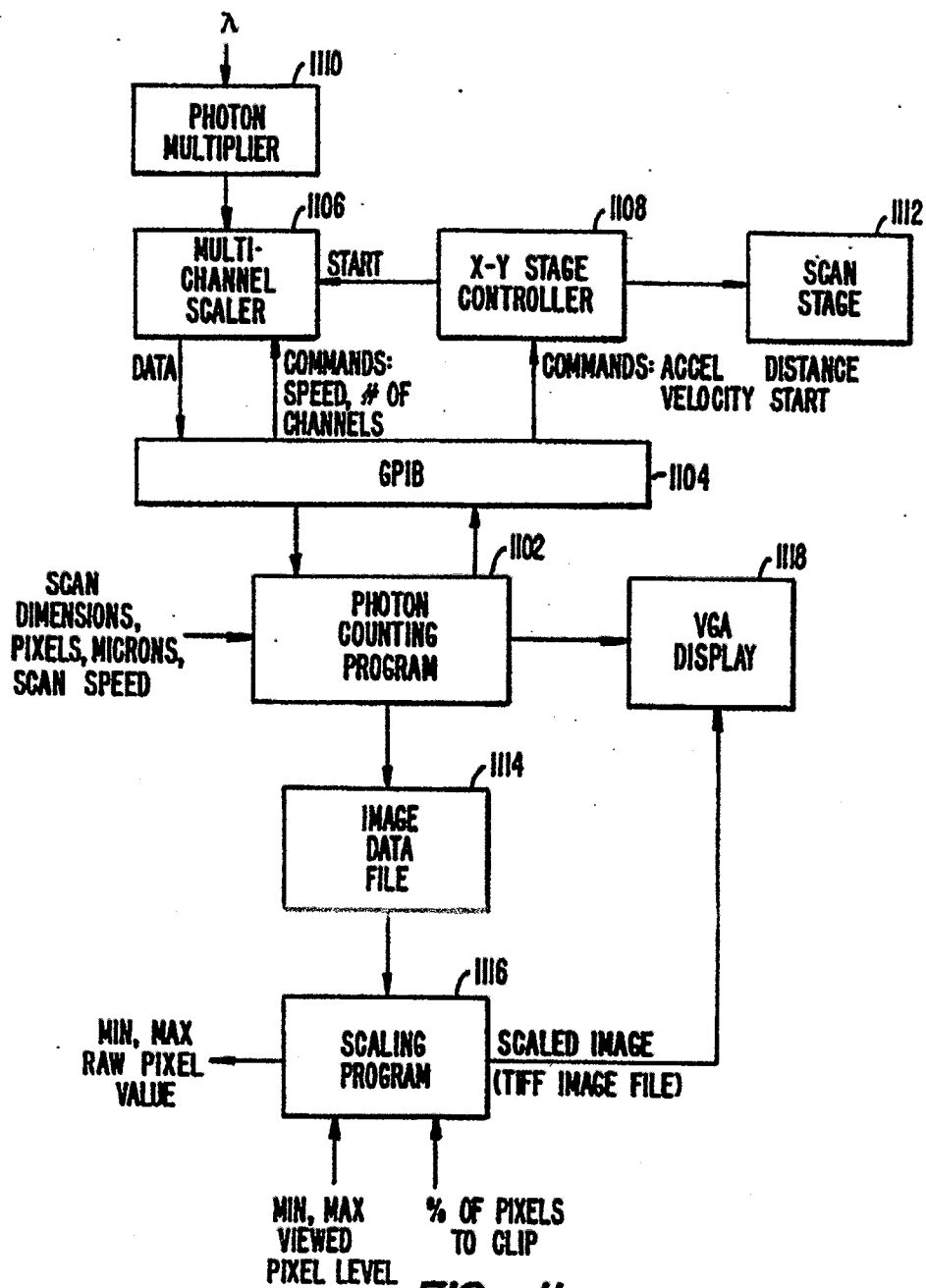


FIG. II.

SUBSTITUTE SHEET